

INTERFACE

royalsocietypublishing.org/journal/rsif

Research



Cite this article: Aliee M, Rock KS, Keeling MJ. 2020 Estimating the distribution of time to extinction of infectious diseases in mean-field approaches. *J. R. Soc. Interface* **17**: 20200540.
<https://doi.org/10.1098/rsif.2020.0540>

Received: 8 July 2020

Accepted: 12 November 2020

Subject Category:

Life Sciences—Mathematics interface

Subject Areas:

biomathematics, biocomplexity, biophysics

Keywords:

birth–death model, sleeping sickness, stochastic infection model, deterministic threshold, disease extinction

Author for correspondence:

Maryam Aliee

e-mail: maryam.aliee@warwick.ac.uk

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.5228093>.

Estimating the distribution of time to extinction of infectious diseases in mean-field approaches

Maryam Aliee^{1,2}, Kat S. Rock^{1,2} and Matt J. Keeling^{1,2}

¹Mathematics Institute, and ²Zeeman Institute for Systems Biology and Infectious Disease Epidemiology Research, University of Warwick, Coventry CV4 7AL, UK

MA, 0000-0003-2268-2295; KSR, 0000-0001-8414-7648; MJK, 0000-0003-4639-4765

A key challenge for many infectious diseases is to predict the time to extinction under specific interventions. In general, this question requires the use of stochastic models which recognize the inherent individual-based, chance-driven nature of the dynamics; yet stochastic models are inherently computationally expensive, especially when parameter uncertainty also needs to be incorporated. Deterministic models are often used for prediction as they are more tractable; however, their inability to precisely reach zero infections makes forecasting extinction times problematic. Here, we study the extinction problem in deterministic models with the help of an effective ‘birth–death’ description of infection and recovery processes. We present a practical method to estimate the distribution, and therefore robust means and prediction intervals, of extinction times by calculating their different moments within the birth–death framework. We show that these predictions agree very well with the results of stochastic models by analysing the simplified susceptible–infected–susceptible (SIS) dynamics as well as studying an example of more complex and realistic dynamics accounting for the infection and control of African sleeping sickness (*Trypanosoma brucei gambiense*).

1. Introduction

For many infectious diseases, the eventual aim of control measures is eradication—completely removing the pathogen from host populations and the environment. This has been accomplished for just two infections—smallpox and rinderpest—but is the target for a number of other diseases, including polio, Guinea worm and yaws [1–3]. A key question is, therefore, predicting the time to eradication, which has clear implications for the likely duration and hence cost of any control programme.

Markov chain models are often used to study the spread of infectious diseases in a population [4,5]. In such approaches, the whole population is partitioned into compartments and the system dynamics are described by the rates of exchange between compartments. Both stochastic and deterministic versions of such models have been developed to study the dynamics of a range of infectious diseases. The results of the deterministic and stochastic versions of the same model generally agree well for large compartment populations, but start to deviate either when there are low numbers of infections or for small population sizes [4,6–9].

Deterministic models provide a mean-field approximation of infection dynamics described by ordinary differential equations (ODEs). The steady states and time evolution of these models are tractable and typically well behaved in the whole phase space. Various analytical and numerical methods are available to solve a set of ODEs; this makes deterministic models computationally efficient and therefore allows the exploration of large regions of parameter space relatively quickly. This is particularly important for fitting such models to available data in order to estimate the underlying mechanistic parameters.

Deterministic descriptions are not very practical, however, when it comes to studying rare events or very small populations. Importantly, the endpoint of an infection (when infection is eradicated) is by definition ambiguous in these models since the populations are represented by continuous variables that never reach zero. In practice, a proxy threshold is needed to determine when extinction may occur; however, it is not clear how such a threshold should be defined. Infections dropping below 1 may seem a natural measure, but in the literature different threshold values are applied [10,11], generating shifts in the predicted extinction times.

In such circumstances, stochastic models are more favourable. In these models, transitions between compartments are generated by stochastic events, the average of which is well captured by the mean-field image but that express divergent behaviours individually. These events and accordingly system state are specified by probability functions allowing for uncommon incidents. Stochastic models are generally computationally expensive; however, they come with a good deal of flexibility and capabilities that we would like to take advantage of to understand the extinction of diseases.

In this article, we will develop a basic framework to study the peri-elimination phase for an infectious disease using a simplified susceptible–infected–susceptible (SIS) model that takes into account susceptible and infected compartments. We use solutions of the corresponding forward Kolmogorov equation to predict the point of infection extinction in deterministic solutions. We then extend this framework to estimate the extinction time of more complicated infection dynamics. While other studies have indeed addressed similar questions around the computation of mean extinction times of epidemiological or other physical systems [12–20], the present study focuses on the computation of the distribution in predicted extinction times, which has not been done elsewhere. This aspect is crucially important as the policy implications of extinction occurring within shorter or longer periods could lead to different decision-making. Indeed, uncertainty in model prediction is not accounted for often enough in epidemiological modelling studies [21]. Representation of uncertainty allows the modeller to answer policy-driven questions such as ‘At what time would we expect there to be a greater than 90% likelihood that elimination would be met?’ or ‘What is the earliest we could hope to achieve zero infections?’, which the mean value does not address.

For this study, we are motivated by a desire to analyse the endpoint of African sleeping sickness (*gambiense* human African trypanosomiasis; gHAT), which is a vector-borne disease transmitted to humans by tsetse [22–25]. Humans transition through different stages of the disease and can either be diagnosed and recover through treatment or typically die without it. Medical intervention campaigns have had great success in bringing down the burden of this disease—from 37 385 in the most recent peak in 1998 to a record low of 953 cases in 2018 [26]. To continue the drive forwards, the World Health Organization (WHO) has a goal to achieve global elimination of transmission (EOT) of sleeping sickness by 2030 [27]. Quantitative methods provide a way to forecast the expected time to meet the goal under a variety of strategies; however, there is a strong need for refinement of the methodology to analyse the distribution of extinction times in a more rigorous mathematical framework. Previous studies have used different threshold criteria for the endpoint for gHAT by

considering when the incidence of new infections falls below one of: 1 per 10 000, 1 per 100 000, less than 0.5 new infections across the entire health zone (approx. 100 000) per year, even to a more severe threshold of less than 1 per 1 000 000 [23,28–30].

2. Model structure

The simple SIS model is a one-dimensional characterization of an infectious disease, in which recovery (by treatment or otherwise) quickly leaves the individual susceptible to further infection [4,5]. For a population of size N , the number of infected individuals can be expressed as

$$\frac{dI}{dt} = \beta SI/N - \gamma I, \quad \text{where } S = N - I. \quad (2.1)$$

Here, β is the transmission rate, γ is the recovery rate and S and I are the numbers of susceptible and infectious individuals, respectively. Another important metric, $R_0 = \beta/\gamma$, is the basic reproduction number. When $R_0 > 1$ this model exhibits sigmoidal growth to the non-zero equilibrium ($I^* = 1 - 1/R_0$), whereas when $R_0 < 1$ we observe exponential decay of infection to zero. Through changes to the transmission or recovery rates, it is possible to modify the *effective reproduction number*, R_{eff} , and push it below this critical threshold. It is this latter case that is of greatest interest here, as we are most concerned with diseases that are bound for eventual extinction (assuming controls remain in place and R_{eff} continues to be less than 1).

The nonlinear SI term in equation (2.1) precludes many analytical results; however, if we are purely interested in extinction it is acceptable to assume that the number of infected individuals is relatively small and hence that $S \approx N$. In this limit, we obtain a simple birth–death process for infections, where birth corresponds to infecting someone and death refers to recovery [31]

$$\frac{dI}{dt} = \beta I - \gamma I = bI - dI. \quad (2.2)$$

Here, β plays the role of the *per capita* birth rate (b) and γ is equivalent to the *per capita* death rate (d). Throughout the rest of this paper, we will use the parameters b and d to stress that we are approximating the dynamics as a birth–death process. Our condition that R_{eff} now corresponds to $d > b$.

2.1. Dynamics without births

We first consider the simplest case where $b = \beta = 0$, and so each infected individual recovers (or is treated) at a Poisson rate d , but no new infections occur. Starting with n infected individuals, the expected time for the first individual to recover is $1/(dn)$. After this recovery event, the population then has $n - 1$ infected individuals and the process can be iterated giving an expected time to extinction of

$$\tau_n = \frac{1}{d} \sum_{m=1}^n \frac{1}{m}. \quad (2.3)$$

Figure 1*a* illustrates how the average extinction time τ_n depends on the initial number of infected people n , showing a logarithmic scaling for large n .

Starting with $I(0) = n$, the deterministic model predicts an exponential decay of the size of infected population

$$I(t) = n \exp(-dt);$$

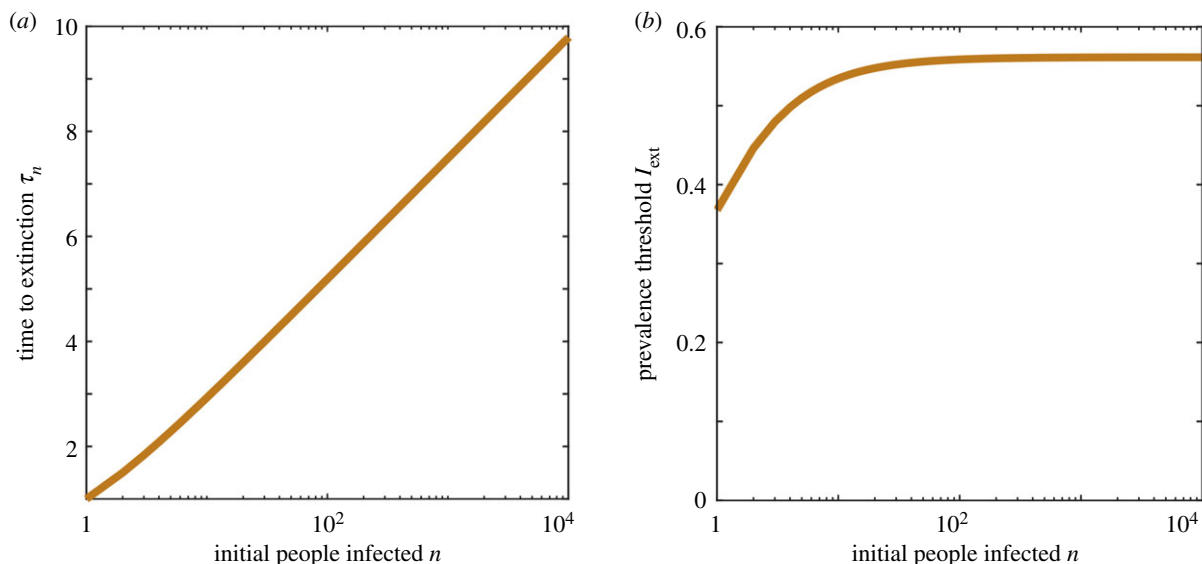


Figure 1. Extinction in a death process model. (a) The average extinction time τ_n and (b) the extinction threshold of infected population I_{ext} are plotted as a function of the initial number of infected population. Here, $d = 1$ per unit time, with no specific units considered.

hence, at the expected extinction time we have

$$\begin{aligned} I_{\text{ext}} &= n \exp(-d\tau_n) \\ &= \exp\left(\log(n) - \sum_{m=1}^n \frac{1}{m}\right), \end{aligned} \quad (2.4)$$

which we postulate gives us a threshold value of the deterministic model that equates to the mean extinction time of the associated stochastic model. Figure 1b shows how the threshold value changes with the initial number of people infected. For large n , the threshold plateaus to $\exp(-E)$, where E is Euler's constant (0.5772), and hence the threshold is 0.5615. Therefore, we observe the useful property that the threshold is largely independent of n .

2.2. Birth–death dynamics

Moving to the situation where b is non-zero, the dynamics are more complex. The expected time to extinction has to be formulated by considering the impact of birth and death from a given number of infected individuals (n). This generates an iterated set of equations for the extinction time starting with n infections

$$\tau_n = \frac{1}{n(d+b)} + \frac{d}{d+b}\tau_{n-1} + \frac{b}{d+b}\tau_{n+1}, \quad (2.5)$$

where the first term is the expected time until an event occurs. The other two terms take into account the possibility of a birth or death event transitioning to $n+1$ or $n-1$ infected individuals. There is a large relevant literature using similar approaches in epidemiology or in other contexts which lead to comparable equations for the mean extinction time [12–17,32]. Equation (2.5) can be expressed as

$$(d+b)\tau_n - d\tau_{n-1} - b\tau_{n+1} = \frac{1}{n}, \quad (2.6)$$

which we solve as a matrix operation ($\mathbf{Ax} = \mathbf{c}$) for $\mathbf{x} = (\tau_0 \dots \tau_M)$ with the boundary conditions $\tau_0 = 0$ and $\tau_{M+1} = \tau_M + 1/(M(d-b))$ for large M , the maximum number of initial people infected.

We note that without loss of generality we can assume that $d = 1$ as this simply acts as a rescaling of time. Figure 2a shows the average extinction time calculated for different choices of b . The average extinction time scales with $1/(d-b)$ and increases logarithmically with n for large values.

With the same approach as in the previous section, we can calculate the threshold of number of people infected $I_{\text{ext}} = n \exp((b-d)\tau_n)$, using the predicted mean extinction time from equation (2.5). Figure 2b shows how the threshold depends on the initial size of the infected population, calculated for different choices of birth rate b . Our results confirm the useful property that the threshold is largely independent of n for moderate numbers of initially infected people (e.g. $n > 500$) and that a scaled threshold (by a factor of $1 - b/d$) reaches the constant value of $\exp(-E)$ (figure 2c).

2.3. Distribution of extinction times

The above approach to formulate the mean extinction times can be extended to higher cumulative moments to estimate the variation and distribution of extinction time in the birth–death model. The second cumulative moment, $S_n^{(2)}$, defined as the expected value of the square of the time to extinction, obeys

$$\begin{aligned} S_n^{(2)} &= \frac{d}{d+b} \left[S_{n-1}^{(2)} + \frac{2}{n(d+b)} \tau_{n-1} + \frac{2}{n^2(d+b)^2} \right] \\ &\quad + \frac{b}{d+b} \left[S_{n+1}^{(2)} + \frac{2}{n(d+b)} \tau_{n+1} + \frac{2}{n^2(d+b)^2} \right]. \end{aligned}$$

Rearranging the above form, and using our previous expression (equation (2.5)), we obtain the simplified sequential relation

$$(d+b)S_n^{(2)} - dS_{n-1}^{(2)} - bS_{n+1}^{(2)} = \frac{2}{n}\tau_n.$$

This can be solved with a similar matrix operation as before ($\mathbf{Ax} = \mathbf{c}$) with the boundary conditions $S_0 = 0$ and $S_{M+1} = S_M + \frac{2}{M(d-b)}\tau_M$ for large M to obtain $S_n^{(2)}$ and hence the variance is given by $V_n = S_n^{(2)} - \tau_n^2$.

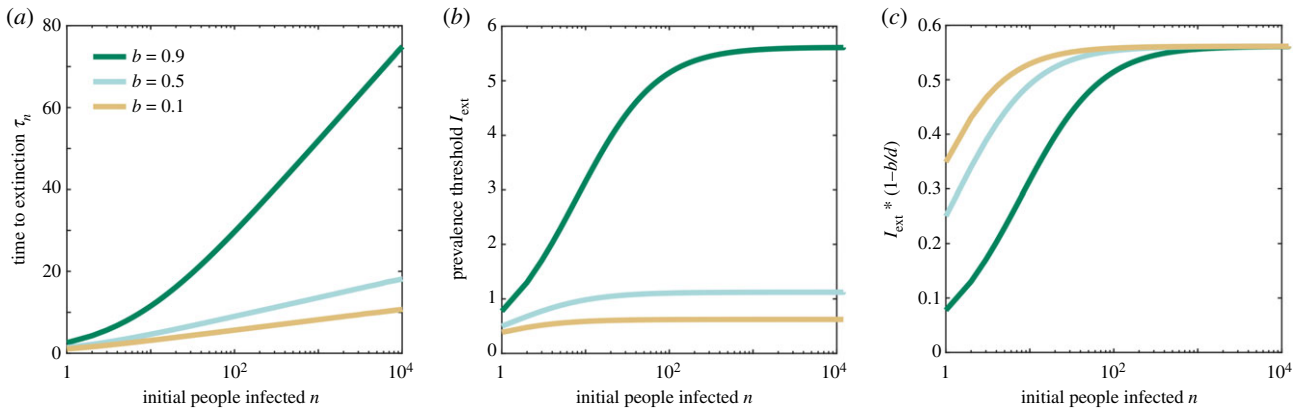


Figure 2. Extinction in a death–birth process model. (a) The average extinction time, τ_n , (b) the extinction threshold of the infected population, I_{ext} , and (c) a rescaled version of this threshold are plotted as a function of the initial number of infected population for three choices of birth rate b . Again, $d = 1$ per unit time.

Knowing that the distribution of extinction times is highly asymmetric with a long right-hand tail, higher cumulative moments help provide a more accurate estimate of prediction intervals and the full distribution of extinction times. We find that the average cube of extinction times, which is used to calculate the third cumulative moment, obeys

$$S_n^{(3)} = \frac{d}{d+b} \left[S_{n-1}^{(3)} + \frac{3}{n(d+b)} S_{n-1}^{(2)} + \frac{6}{n^2(d+b)^2} \tau_{n-1} + \frac{6}{n^3(d+b)^3} \right] + \frac{b}{d+b} \left[S_{n+1}^{(3)} + \frac{3}{n(d+b)} S_{n+1}^{(2)} + \frac{6}{n^2(d+b)^2} \tau_{n+1} + \frac{6}{n^3(d+b)^3} \right],$$

which can be simplified as

$$(d+b)S_n^{(3)} - dS_{n-1}^{(3)} - bS_{n+1}^{(3)} = \frac{3}{n} S_n^{(2)}.$$

This approach can be extended to calculate sequential relations of cumulative moments of higher orders m

$$(d+b)S_n^{(m)} - dS_{n-1}^{(m)} - bS_{n+1}^{(m)} = \frac{m}{n} S_n^{(m-1)}.$$

The full set of cumulative moments allows the probability distribution of extinction times to be calculated. Our numerical solutions suggest that the first four moments can give a good estimate of the distribution, approximated by a generalized F probability distribution function with four free parameters. This allows us to determine the prediction intervals for a given set of parameters (for details, see the electronic supplementary material).

As an alternative approach, we consider the Kolmogorov forward equations to describe the time evolution of probabilities for the numbers infected [33]. These high-dimensional models are a numerically exact realization of the full stochastic model, which can be computed for the SIS model because of the relatively low number of states. The Kolmogorov forward equations are given by

$$\frac{dP_i}{dt} = -i(b+d)P_i + d(i+1)P_{i+1} - b(i-1)P_{i-1}.$$

Here, P_i depicts the probability of having i individuals infected. The solution of this equation is given by $\mathbf{P} = \exp(\mathbf{A}t) \cdot \mathbf{P}_0$ for the vector of P_i s over time, where \mathbf{A} is the coefficient matrix and \mathbf{P}_0 is the initial state of the system. We can accordingly calculate the probability of extinction over time and therefore the distribution.

Figure 3 compares the estimated F distributions of extinction times from the cumulative moments with the corresponding prediction intervals with the solutions of the Kolmogorov forward equations. Again, without loss of generality we have taken $d = 1$, and varied $b < d$. The distribution shifts to the right and flattens as birth rate increases. These results demonstrate the very good agreement between the distributions, as well as the mean values and 95% prediction intervals. This analysis, therefore, provides confidence that our simple approach is a useful tool to estimate extinction times and prediction intervals for any choice of b and d . In the electronic supplementary material, we plot how variance, skewness and kurtosis change as a function initial population infected n .

3. Application to the complex dynamics of sleeping sickness

We now investigate how well our theory holds against more realistic, and hence more complex, simulation models. Our previously developed model for gHAT [28,34] is high dimensional, describing the dynamics of vectors and two human risk groups. The model accounts for the natural history of disease progression: infected humans initially develop mild disease (stage 1) before progressing to more severe disease (stage 2) with neurological symptoms as the parasite crosses the blood–brain barrier. The infectious period ends with either death or successful diagnosis and treatment by screening programmes [35]. This model is fitted to the human case incidence data recorded in different administrative regions (health zones of approximately 100 000 people) across the Democratic Republic of Congo to estimate the free parameters of the deterministic model [34].

Using a Markov chain Monte Carlo approach, posterior parameter sets that match available case data have previously been generated (see [34] and the electronic supplementary material for details). The deterministic equations can be numerically solved to generate the disease dynamics including the number of humans infected and new transmissions each year for posterior parameter sets. We use these results and our knowledge of the underlying epidemiology to determine appropriate values of d and b that approximate the linear birth–death process (see electronic supplementary material for details). This allows us to generate the first

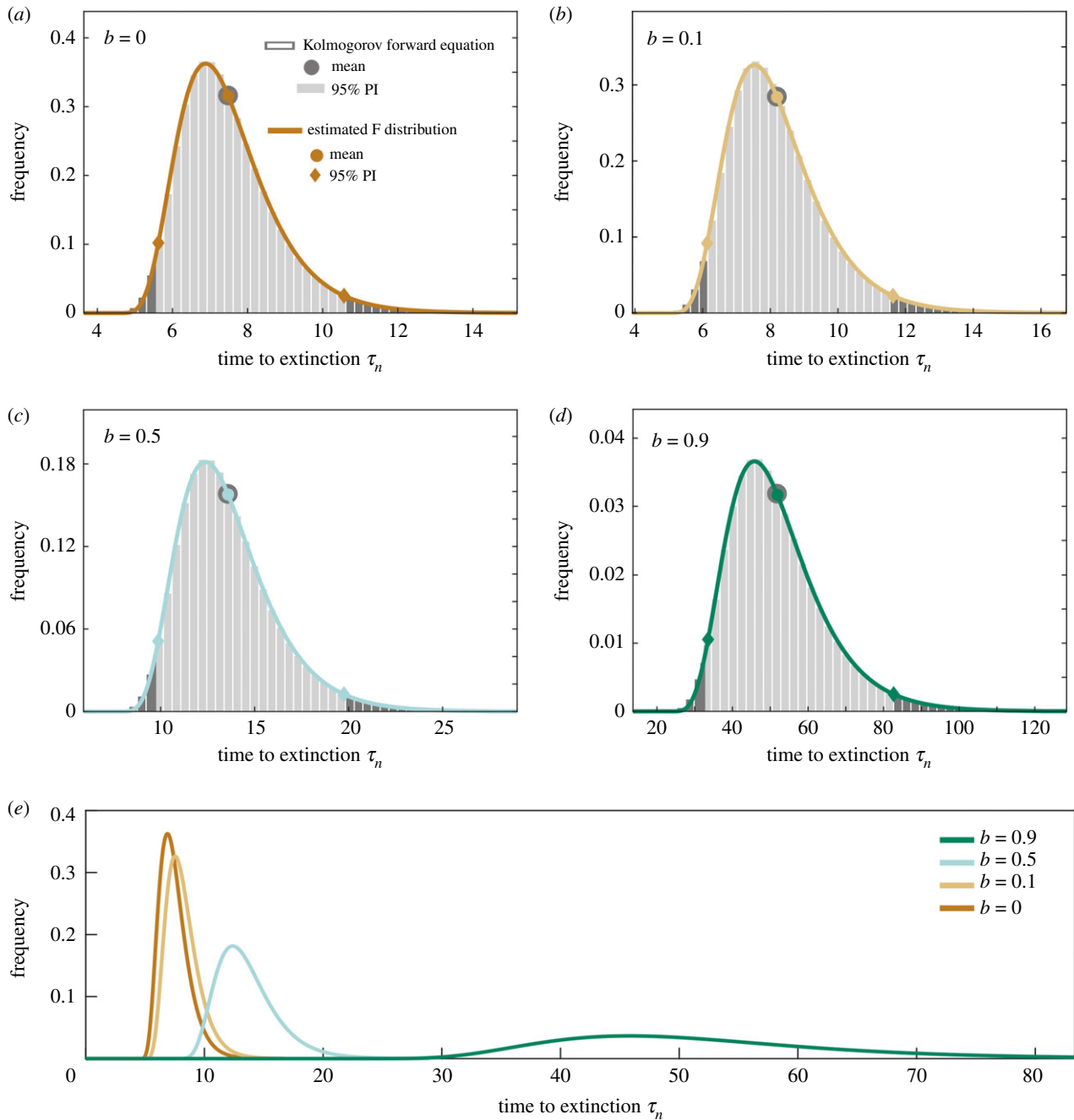


Figure 3. Probability distribution of extinction time τ_n in the model of the birth–death process starting with $n = 1000$. (a–d) Grey bars represent the solutions of the Kolmogorov forward equation (light grey specifies 95% prediction interval (PI)). The coloured line shows the F distribution estimated from the first four moments of the birth–death process. Corresponding mean values and 95% PIs are indicated for each dataset. Each panel corresponds to a birth rate b , when $d = 1$ per unit time for all. (e) Comparing the estimated F distribution functions plotted for all values of birth rate b .

four moments, and accordingly the distribution of extinction times for specific d and b rates following the framework developed in the previous sections.

To verify these results, we run the stochastic discrete simulations of the gHAT model using the tau-leaping approach described elsewhere [25,36] and select a 1 day time step which approximates the full dynamics well due to the long time scale of gHAT infections (a Gillespie method could be used for more accurate simulations; however, it is unlikely to modify the results we present here). We compute the extinction times for multiple stochastic realizations. These simulations assume that active screening (the fundamental driver of case identification and control) continues at a constant level, leading to decaying infection levels (electronic supplementary

material, figures). Figure 4 compares the F probability distribution function derived from the estimated cumulants with the distribution of extinction times of 1 000 000 stochastic simulations for two different parameter sets. The parameter sets represent the most likely of all posterior parameter sets found by fitting to data for two health zones named Mosango (a ‘moderate-risk’ zone) and Kwamouth (a ‘high-risk’ zone) [34] that reveal different infection dynamics over time, and therefore different predicted extinction times. Despite the complexity of the model for gHAT dynamics, there is still a very good agreement between the birth–death model with F distribution predictions and full stochastic gHAT infection dynamics (with differences in mean and 95% prediction interval (PI) of less than a year), confirming that the analysis based

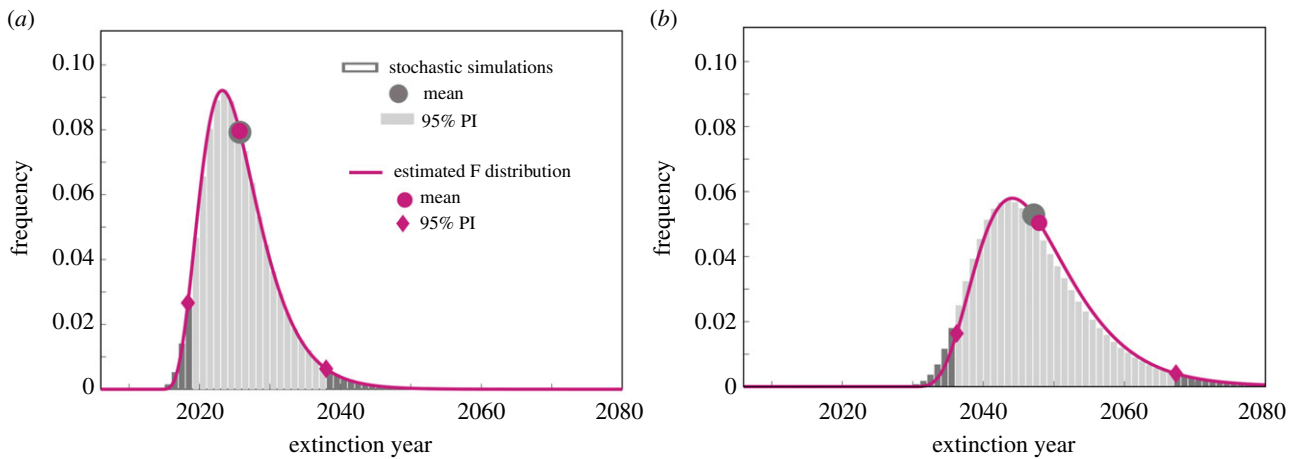


Figure 4. Probability distribution of extinction time for the full (17-dimensional) gHAT model for two health zones, (a) Mosango and (b) Kwamouth with 30% active screening. Grey bars represent the solutions of one million stochastic simulations (light grey specifies 95% PI). The purple line shows the F distribution estimated from the first four moments of the simplified birth–death process approach ($b = 0.42$, $d = 0.67$ for Mosango and $b = 0.48$, $d = 0.64$ for Kwamouth). Corresponding mean values and 95% prediction intervals are indicated for each dataset.

on the birth–death model can deliver robust and accurate estimates of mean extinction time and the prediction intervals. Examples using other levels of screening coverage are shown in the electronic supplementary material. In these examples, for a fixed parameter set, we computed that the deterministic threshold, I_{ext} , to compute the expected extinction time was 1.50 and 2.26 for Mosango and Kwamouth, respectively. Other threshold values would need to be calculated for alternative parameter sets and alternative screening strategies for the same health zones.

It should be noted that this study provides a simplified picture of eradication of a disease taking into account instant feedback between transmission and the infected population. However, in epidemiology, other measures such as elimination of transmission may be of interest. In the case of complex dynamics like gHAT, because of the phase difference of humans and vectors as well as the slow progression of the disease and recovery process, there is a delay between elimination of transmission and disease extinction. We speculate that the lag period can be estimated based on the underlying dynamics.

Another challenge is to extend such an analysis to allow for more general noise models, while our current stochastic models mainly account for Poisson events. In the case of gHAT dynamics, we are aware of possible over-dispersion terms that can be considered in the infection dynamics [25,34]. It is important to study how the predicted distribution of extinction time will be modified by such noise models.

4. Conclusion

In this article, we study the extinction times of diseases when $R_{eff} < 1$ starting from a finite population infected, using results from a simplified birth–death model. We formulate the mean extinction time that is in agreement with the previous studies in the analogous limits [13–17]. Beyond that, our analysis provides a novel method to approximate the distribution of extinction times by calculating additional moments of extinction time. Our results agree well with the solutions of forward Kolmogorov equations, which offer a numerical exact prediction of the birth–death dynamics.

Our findings are a vital extension to deterministic models, allowing a robust and rigorous assessment of extinction in a modelling framework that cannot naturally capture this phenomenon. We considered the simplistic assumption of constant birth and death rates; however, the results can be easily extended by modifying the coefficient matrix when these rates are functions of the infected population.

Moreover, our analysis helps us to quantify a simple threshold that can be used to determine the mean extinction time from deterministic solutions. We showed that this threshold asymptotes to a universal plateau given by the effective reproduction ratio for large initial numbers of infected individuals. This asymptotic threshold is likely to be a reasonable approximation for large populations where the number of infections is also likely to be large.

Interestingly, the birth–death approach can be applied to diseases with more complex disease dynamics in order to predict their extinction times. We showed that it is valid for the example of local elimination of sleeping sickness, yet in a broad sense the dynamics are approximately SIS decaying to zero infection. Our study can be extended for SIR models, where the system state is two-dimensional by considering the recovered (R) compartment. A non-trivial modification of the birth–death process would be helpful to consider to understand the richer dynamics of the SIR model [37].

Data accessibility. This work does not include any original data. To describe the gHAT dynamics, we used the model and the fitted parameters from a published article that is cited in the main text and discussed in detail in the electronic supplementary material.

Authors' contributions. All authors contributed to developing the model and writing the manuscript.

Competing interests. We declare we have no competing interests.

Funding. This work was supported by the Bill and Melinda Gates Foundation (www.gatesfoundation.org) through the NTD Modelling Consortium (OPP1184344) (M.A., K.S.R. and M.J.K.) and the Human African Trypanosomiasis Modelling and Economic Predictions for Policy (HAT MEPP) project (OPP1177824) (K.S.R. and M.J.K.). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements. The authors thank Dr Erick Mwamba Miaka and PNLTHA for original data collection, and the World Health Organization (WHO) for data access (in the framework of the WHO HAT Atlas [27]).

1. Hinman AR. 1998 Global progress in infectious disease control. *Vaccine* **16**, 1116–1121. (doi:10.1016/S0264-410X(98)80107-2)
2. Hopkins DR. 2013 Disease eradication. *N. Engl. J. Med.* **368**, 54–63. (doi:10.1056/NEJMra1200391)
3. Nieddu GT, Billings L, Kaufman JH, Forgoston E, Bianco S. 2017 Extinction pathways and outbreak vulnerability in a stochastic Ebola model. *J. R. Soc. Interface* **14**, 20160847. (doi:10.1098/rsif.2016.0847)
4. Keeling MJ, Rohani P. 2011 *Modeling infectious diseases in humans and animals*. Princeton, NJ: Princeton University Press.
5. Anderson RM, May RM. 1991 *Infectious diseases of humans*. Oxford, UK: Oxford Science Publications.
6. Näsell I. 2002 Stochastic models of some endemic infections. *Math. Biosci.* **179**, 1–19. (doi:10.1016/S0025-5564(02)00098-6)
7. Keeling MJ, Ross JV. 2008 On methods for studying stochastic disease dynamics. *J. R. Soc. Interface* **5**, 171–181. (doi:10.1098/rsif.2007.1106)
8. Keeling MJ, Wilson HB, Pacala SW. 2000 Reinterpreting space, time lags, and functional responses in ecological models. *Science* **290**, 1758–1761. (doi:10.1126/science.290.5497.1758)
9. Keeling MJ. 2000 Simple stochastic models and their power-law type behaviour. *Theor. Popul. Biol.* **58**, 21–31. (doi:10.1006/tpbi.2000.1475)
10. Hollingsworth TD *et al.* 2015 Quantitative analyses and modelling to support achievement of the 2020 goals for nine neglected tropical diseases. *Parasites Vectors* **8**, 630. (doi:10.1186/s13071-015-1235-1)
11. Truscott JE, Werkman M, Wright JE, Farrell SH, Sarkar R, Ásbjörnsdóttir K, Anderson RM. 2017 Identifying optimal threshold statistics for elimination of hookworm using a stochastic simulation model. *Parasites Vectors* **10**, 321.
12. Ushio Sumita YM. 1985 On first passage time structure of random walks. *Stoch. Process. Appl.* **20**, 133–147. (doi:10.1016/0304-4149(85)90021-3)
13. Näsell I 2011 *Extinction and quasi-stationarity in the stochastic logistic SIS model*. Lecture Notes in Mathematics, vol. 2022. Berlin, Germany: Springer.
14. Britton T, Neal P. 2010 The time to extinction for a stochastic SIS-household-epidemic model. *J. Math. Biol.* **61**, 763–769. (doi:10.1007/s00285-009-0320-5)
15. Clancy D, Tjia E. 2018 Approximating time to extinction for endemic infection models. *Methodol. Comput. Appl. Probab.* **20**, 1043–1067. (doi:10.1007/s11009-018-9621-8)
16. Andersson H, Britton T. 2000 Stochastic epidemics in dynamic populations: quasi-stationarity and extinction. *J. Math. Biol.* **41**, 559–580. (doi:10.1007/s002850000060)
17. Doering CR, Sargsyan KV, Sander LM. 2005 Extinction times for birth-death processes: exact results, continuum asymptotics, and the failure of the Fokker-Planck approximation. *Multiscale Model. Simul.* **3**, 283–299. (doi:10.1137/030602800)
18. Dykman MI, Schwartz IB, Landsman AS. 2008 Disease extinction in the presence of random vaccination. *Phys. Rev. Lett.* **101**, 078101. (doi:10.1103/PhysRevLett.101.078101)
19. Forgoston E, Moore RO. 2018 A primer on noise-induced transitions in applied dynamical systems. *SIAM Rev.* **60**, 969–1009. (doi:10.1137/17M1142028)
20. Assaf M, Meerson B. 2010 Extinction of metastable stochastic populations. *Phys. Rev. E* **81**, 021116. (doi:10.1103/PhysRevE.81.021116)
21. Gilbert JA, Ance Meyers L, Galvani AP, Townsend JP. 2014 Probabilistic uncertainty analysis of epidemiological modeling to guide public health intervention policy. *Epidemics* **6**, 37–45. (doi:10.1016/j.epidem.2013.11.002)
22. Rock KS, Torr SJ, Lumbala C, Keeling MJ. 2017 Predicting the impact of intervention strategies for sleeping sickness in two high-endemicity health zones of the Democratic Republic of Congo. *PLoS Negl. Trop. Dis.* **11**, e0005162. (doi:10.1371/journal.pntd.0005162)
23. Rock KS, Pandey A, Ndeffo-Mbah ML, Atkins KE, Lumbala C, Galvani A, Keeling MJ. 2017 Data-driven models to predict the elimination of sleeping sickness in former Equateur province of DRC. *Epidemics* **18**, 101–112. (doi:10.1016/j.epidem.2017.01.006)
24. Rock KS *et al.* 2018 Assessing strategies against gambiense sleeping sickness through mathematical modeling. *Clin. Infect. Dis.* **66**(suppl. 4), S286–S292. (doi:10.1093/cid/ciy018)
25. Castaño MS, Aliee M, Mwamba Miaka E, Keeling MJ, Chitnis N, Rock KS. 2020 Screening strategies for a sustainable endpoint for gambiense sleeping sickness. *J. Infect. Dis.* **221**, S539–S545. (doi:10.1093/infdis/jiz588)
26. World Health Organization. 2019 Global health observatory data. See <https://www.who.int/data/gho/data/indicators/indicator-details/GHO/hat-tb-gambiense>.
27. Franco JR, Cecchi G, Priotto G, Paone M, Diarra A, Grout L, Simarro PP, Zhao W, Argaw D. 2018 Monitoring the elimination of human African trypanosomiasis: update to 2016. *PLoS Negl. Trop. Dis.* **12**, e0006890. (doi:10.1371/journal.pntd.0006890)
28. Rock KS, Torr SJ, Lumbala C, Keeling MJ. 2015 Quantitative evaluation of the strategy to eliminate human African trypanosomiasis in the DRC. *Parasites Vectors* **8**, 532. (doi:10.1186/s13071-015-1131-8)
29. Castaño MS *et al.* 2019 Assessing the impact of data aggregation in model predictions of HAT transmission and control activities. (<https://www.medrxiv.org/content/10.1101/19005991v2>)
30. Huang C, Crump RE, Brown P, Spencer SEF, Miaka EM, Chancy S, Keeling MJ, Rock KS. 2020 Shrinking the gHAT map: identifying target regions for enhanced control of gambiense human African trypanosomiasis in the Democratic Republic of Congo. (<https://www.medrxiv.org/content/10.1101/2020.07.03.20145847v1>)
31. Britton T. 2010 Stochastic epidemic models: a survey. *Math. Biosci.* **225**, 24–35. (doi:10.1016/j.mbs.2010.01.006)
32. Sumita U. 1984 On conditional passage time structure of birth-death processes. *J. Appl. Probab.* **21**, 10–21. (doi:10.2307/3213660)
33. Kolmogorov AN. 1993 *Foundations of the theory of probability*. New York, NY: Chelsea Publishing.
34. Crump RE, Huang C-I, Knock E, Spencer SEF, Brown PE, Mwamba Miaka E, Chancy S, Keeling MJ, Rock KS. 2020 Quantifying epidemiological drivers of gambiense human African trypanosomiasis across the Democratic Republic of Congo. (<https://www.medrxiv.org/content/10.1101/2020.06.23.20138065v2>)
35. World Health Organization. 2013 Control and surveillance of human African trypanosomiasis: report of a who expert committee. Technical report. See <https://apps.who.int/iris/handle/10665/95732>.
36. Davis CN, Rock KS, Mwamba Miaka E, Keeling MJ. 2019 Village-scale persistence and elimination of gambiense human African trypanosomiasis. *PLoS Negl. Trop. Dis.* **13**, e0007838. (doi:10.1371/journal.pntd.0007838)
37. Keeling MJ, Grenfell BT. 2002 Understanding the persistence of measles: reconciling theory, simulation and observation. *Proc. R. Soc. Lond. B* **269**, 335–343. (doi:10.1098/rspb.2001.1898)